

ски значимыми. Физическая работоспособность – значимое различие имеет признак толерантность к физической нагрузке, сердечно-сосудистая система – диастолическое и систолическое артериальное давление, липидный обмен - липопротеиды низкой плотности, в группах биохимия крови и гормональный статус получили результаты, что для всех показателей  $p > 0,05$ , следовательно различия между выборками статистически не значимы. Иммунологический статус – значимые циркулирующие иммунные комплексы в сыворотке крови и концентрация иммуноглобулина А в сыворотке крови. В группах состояние калликреин-кининовой системы и окислительная способность плазмы крови для всех показателей  $p < 0,05$ , следовательно различия статистически значимы.

Данные по результатам исследования будут использоваться при формировании базы знаний и создании интеллектуальной системе поддержки принятия врачебных решений, а так же учитываться при сборе необходимых признаков для вновь прибывших пациентов.

#### ЛИТЕРАТУРА

1. Гублер Е.В. Вычислительные методы анализа и распознавания патологических процессов, 1978, 269 с.
2. Голованова И.С. Выбор информативных признаков. Оценка информативности. [Электронный ресурс]. URL: <http://ime.tpu.ru/study/discypliny/INF-PR.pdf> (дата обращения: 10.10.2017).
3. Мастицкий С.Э., Шитиков В.К. Статистический анализ и визуализация данных с помощью R. [Электронный ресурс]. URL: <http://r-analytics.blogspot.ru/> (дата обращения: 21.10.2017).

#### ЦИФРОВАЯ ЭКОНОМИКА: ОБЛАЧНЫЕ ТЕХНОЛОГИИ В ЗАДАЧАХ МАТЕМАТИЧЕСКОГО АНАЛИЗА В СФЕРЕ МЕДИЦИНЫ

*Зими́на Е.Ю.*

*НИУ ВШЭ, [ezimina@hse.ru](mailto:ezimina@hse.ru)*

#### DIGITAL ECONOMY: CLOUD TECHNOLOGIES IN PROBLEMS OF MATHEMATICAL ANALYSIS OF MEDICAL DATA

*Zimina E. U.*

*Higher School of Economics, [ezimina@hse.ru](mailto:ezimina@hse.ru)*

**Abstract.** The article includes the observation of the cloud services and technologies usage in telemedicine. The article contains a review of mathematical analysis of medical data using cloud technology, which produces storage, analysis and forecasting on the basis of obtained data. In addition, the possibility of integrating cloud technologies with external systems is considered.

**Key words:** Telemedicine, digital economy, cardiology, big data

#### Введение

По прогнозу института McKinsey на 2025 год в сфере цифровой экономики наиболее значимыми областями применения технологий будут Mobile Internet, Automation of knowledge work, Internet of Things и Cloud. [1]

На рис 1 приводятся прогнозные оценки мирового рынка по каждой из названных технологий, суммарный объем рынка которых составит около 30 триллионов долларов. При том, что доля медицинского использования этих технологий оценивается ориентировочно в 30%.

Доля нефтегазового сектора мировой экономики оценивается при этом только в 1.5 триллиона долл.

Исходя из изложенного, актуальным представляется применение технологий цифровой экономики в области медицины.

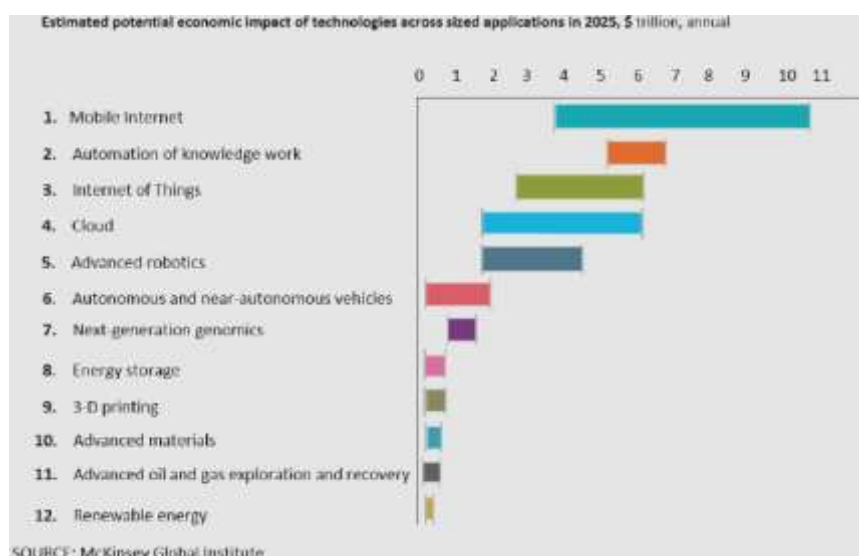


Рисунок 1 – Прогнозные оценки мировых рынков на 2025 год.

Одним из индикаторов роста количества работающих в цифровой экономике является количество публикаций по Большим данным: "сырью" цифровой экономики.

На рисунке 2 представлен график роста количества публикаций, связанных с Большими данными за последние годы. [2]

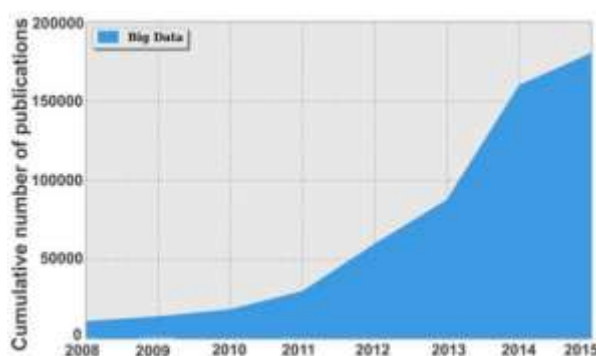


Рисунок 2 – Общее число публикаций со ссылкой на Большие данные.

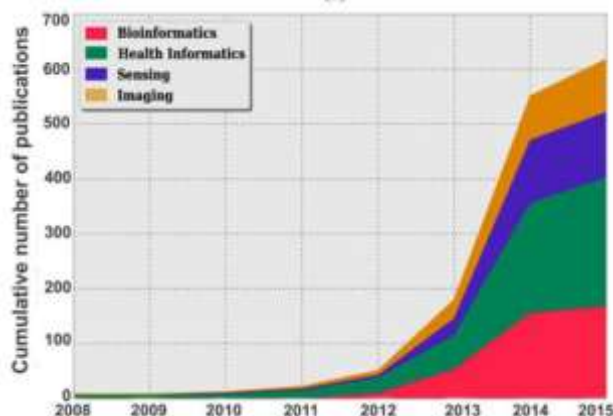


Рисунок 3 – Общее число публикаций по исследованиям в области здравоохранения со ссылкой на Большие данные.

Таким образом, по сведениям академии Google (Google Scholar) на графике рисунка 3 наблюдается резкий рост количества публикаций со ссылкой на Большие данные среди всех публикаций по исследованиям в области здравоохранения. [2]

Отсюда следует, что в настоящее время в медицине широкое использование находит применение информационных технологий, в том числе и Больших данных, что является телемедициной. [3] Актуальность тематики для РФ подчеркивается принятием Федерального Закона от 29 июля 2017 года N 242-ФЗ "О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья", который определяет порядок оказания медпомощи с применением телемедицинских технологий. [3]

Медицина является одной из сфер человеческой деятельности, в которую можно внедрить применение анализа больших данных.

По данным Всемирной Организации Здравоохранения на 2015 год почти 30% причиной всех смертей являются ишемическая болезнь сердца и инсульт. [4] Такая значительная распространённость хронической сердечной недостаточности ведёт к высокой смертности кардиологических больных и требует больших затрат на лечение.

Исследования в направлении телемедицины проводятся в МИЭМ НИУ ВШЭ рамках проекта по созданию централизованной системы ранней массовой диагностики (ЦСРМД) заболеваний населения РФ вне пределов лечебно-профилактических учреждений (ЛПУ), обеспечивающей персонализированный мониторинг, анализ и прогнозирование развития состояния пациента в реальном времени до его первого появления в ЛПУ, а также сопровождение пациентов после постановки диагноза и выписки из ЛПУ.

Системообразующей компонентой ЦСРМД в данном проекте является облако, объединяющее в единую АС с одной стороны оконечные мобильные устройства и группировку обучаемых компьютеров, реализованных в ЦИАС, а с другой стороны - линейный медицинский персонал, принимающий решения по сценариям взаимодействия с пациентами.

Разрабатывается система диагностики заболеваний, обеспечивающая персонализированный мониторинг, анализ и прогнозирование развития состояния пациента в реальном времени.

### **Облачные технологии**

Облако позволяет принимать и обрабатывать кардиограммы (любого числа отведений), фотоплетизмограммы (пульсовая волна), эхокардиограммы, сейсмокардиограммы для дальнейшего проведения расчётов. Во время проведения ЭКГ используется такое понятие, как электрокардиографические отведения (разница потенциалов в электрокардиографии). Для глубокого анализа требуется наличие первого отведения, но любые дополнительные отведения только уточняют дальнейший анализ. [5]

Облако пополняется посредством API и специализированных адаптеров, предназначенных для самостоятельной загрузки данных и адаптеров, предназначенных которые служат для чтения форматов ЭКГ, полученных со специализированных устройств.

Облако предназначено для хранения и обработки с помощью технологий Big Data информации полученной из специальных учреждений или мобильных устройств пользователя, что является собой телемедицину.

Результаты анализа также предоставляются посредством API и специальных адаптеров для интеграции с внешними системами.

Расчёты производятся на основе частотных и энергетических характеристик полученных сигналов. Происходит выделение типовых и переходных частотных состояний для каждого пациента в отдельности и для целевых выборок сигналов, разделённых по возрасту, полу и наличию патологий у пациента.

В ходе анализа определяется близость пациента к одному из определенных частотных состояний по данным одного обследования. Отслеживается движение состояния пациента между типовыми состояниями самого пациента и целевых выборок для определения тенденции развития патологий, действия медикаментов, послеоперационной реабилитации.

Определяется биологический возраст сердечно-сосудистой системы пациента, как отдельный вид траектории. Отслеживаются острые состояния пациента требующих профессиональной помощи.

По удлинению траектории движения состояния пациента между типовыми состояниями самого пациента и целевых выборок производится прогнозирование будущих состояний пациента на ближайшие несколько дней.

При поступлении новой кардиограммы в БД инициируются 2 вида обработок: обработка кардиограммы относительно текущего пациента и обработка обновленного множества всех кардиограмм из БД.

#### **Обработка кардиограмм**

Обработка новой поступившей кардиограммы производится в последовательном выполнении нескольких видов заданий:

- подготовка кардиограммы к дальнейшим расчетам;
- расчёт спектра Фурье и его характеристик (гармоники и энергии) для нефiltroванного сигнала, сигнала с исправленной изолинией и сигнала с исправленной изолинией и очищенного от дефектов;
- расчет формы спектра для сигнала с исправленной изолинией и очищенного от дефектов;
- перерасчет кластеров кардиограмм во всей серии кардиограмм пациента с различными параметрами кластеризации;
- расчет показателей тенденции пациента (биологического возраста).

Обработка обновленного множества всех кардиограмм производится по расписанию и для различных целевых групп по возрасту, полу, информации о патологиях для различных параметров кластеризации.

Возрастные группы выделяются следующие: все возраста, и 5 отдельных возрастных групп от молодых к старым. Группы по полу выделяются следующие: мужчины и женщины.

По патологиям выделяются следующие группы:

- размеченная выборка кардиограмм по диагнозу ишемия (подтвержденная и не подтвержденная)
- размеченная выборка кардиограмм по диагнозу туберкулез (подтвержденная и не подтвержденная)
- размеченная выборка кардиограмм по диагнозу диабет (подтвержденная и не подтвержденная).

Если в группе с момента выполнения последней кластеризации не появилось новых кардиограмм, запуск новой кластеризации не производится.

При обработке отдельной кардиограммы перед проведением расчетов и исследования у неё необходимо исправить изолинию, т.е. удалить дрейф изолинии, и очистить её от сильных дефектов, т.к. это сильно портит частотные характеристики кардиограммы, необходимый необходимый для дальнейшего анализа.

Дрейф изолинии удаляется двухмедианным методом, для этого из исходного сигнала поэлементно вычитается рассчитанная посредством двух медиан изолиния.

#### **Кластеризация выборки кардиограмм**

Кластеризация производится для множества кардиограмм, выделенных по определенным признакам. Для заданий кластеризация производится для всех кардиограмм отдельного пациента с различными параметрами кластеризации по выделенным формам спектра. [6]

Также кластеризация может проводиться на основе формы, полученной другим алгоритмом, на основе усредненного кардиоцикла и на основе амплитуд или спектральных плотностей гармоник. [7]

Для каждой кардиограммы в рамках этого задания рассчитываются эти 6 значений, которые доступны для внешних систем как результаты расчётов посредством API или специализированных адаптеров. [8]

### **Заключение**

В статье рассматриваются возможности использования анализа Больших данных в одной из сфер цифровой экономике – телемедицине. Основным результатом проведенного исследования является реализация обработки и хранения медицинских данных для дальнейшего проведения расчётов в облаке.

В дополнение к этому, проведен обзор возможностей анализа и мониторинга медицинских данных, а также прогнозирования в задачах математического анализа медицинских данных с помощью облачных технологий.

### **ЛИТЕРАТУРА**

1. Disruptive technologies: Advances that will transform life, business, and the global economy [Электронный ресурс] / Институт McKinsey – Режим доступа: <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/disruptive-technologies>. (Дата обращения: 12.09.2017 г.).
2. Big Data for Health [Электронный ресурс] / Javier Andreu, Carmen C. Y. Poon, Robert D. Merrifield, Guang-Zhong Yang – Режим доступа: [https://www.researchgate.net/publication/280124446\\_Big\\_Data\\_for\\_Health](https://www.researchgate.net/publication/280124446_Big_Data_for_Health). (Дата обращения: 01.09.2017 г.).
3. Федеральный Закон от 29 июля 2017 года N 242-ФЗ "О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья" // Собрание законодательства РФ. — 2017. — N 242-ФЗ.
4. 10 ведущих причин смерти в мире [Электронный ресурс] / Всемирная Организация Здравоохранения – Режим доступа: <http://www.who.int/mediacentre/factsheets/fs310/ru/>. (Дата обращения: 13.11.2017 г.).
5. Novopashin M.A., Shmid A.V. D.SC. Berezin A.A. PhD, Forrester's Concept in Modeling Heart Dynamics. // IOSR Journal of Computer Engineering. – May-Jun. 2017. – Vol. 19. – Issue 3, Ver.2.
6. Berezin, A.A. Resonant interaction between the Fermi-Pasta-Ulam recurrences.// Bulletin of the Lebedev Physics Institute. FIAN. – 2004. – N. 3. – P. 13.
7. Fermi E., Pasta J., and Ulam S. In: Collected Papers of E. Fermi. – 1955. – Vol. 2, 978 p.
8. Moyer, V.A. Screening for coronary heart disease with electrocardiography: U.S. Preventive Services Task Force recommendation statement.// Annals of Internal Medicine. . – 2002. – P. 157.

### **ИСПОЛЬЗОВАНИЕ ВИЗУАЛИЗАЦИИ ПРИ АНАЛИЗЕ МЕДИЦИНСКИХ ДАННЫХ**

*И.А. Лызин, научный руководитель О.В.Марухина  
(г. Томск, Томский политехнический университет)  
e-mail: i-lyzin@mail.ru*

### **USE OF VISUALIZATION IN THE ANALYSIS OF MEDICAL DATA**

*I.A. Lyzin, scientific supervisor O.V.Marukhina  
(Tomsk, Tomsk Polytechnic University)*

**Abstract.** The powerful facility for information analysis is visualization facilities. Visualization of data it is a visual representation of the arrays of different information. In the article describe the importance of visualization of data. The programming language for statistical data handling and operation with R